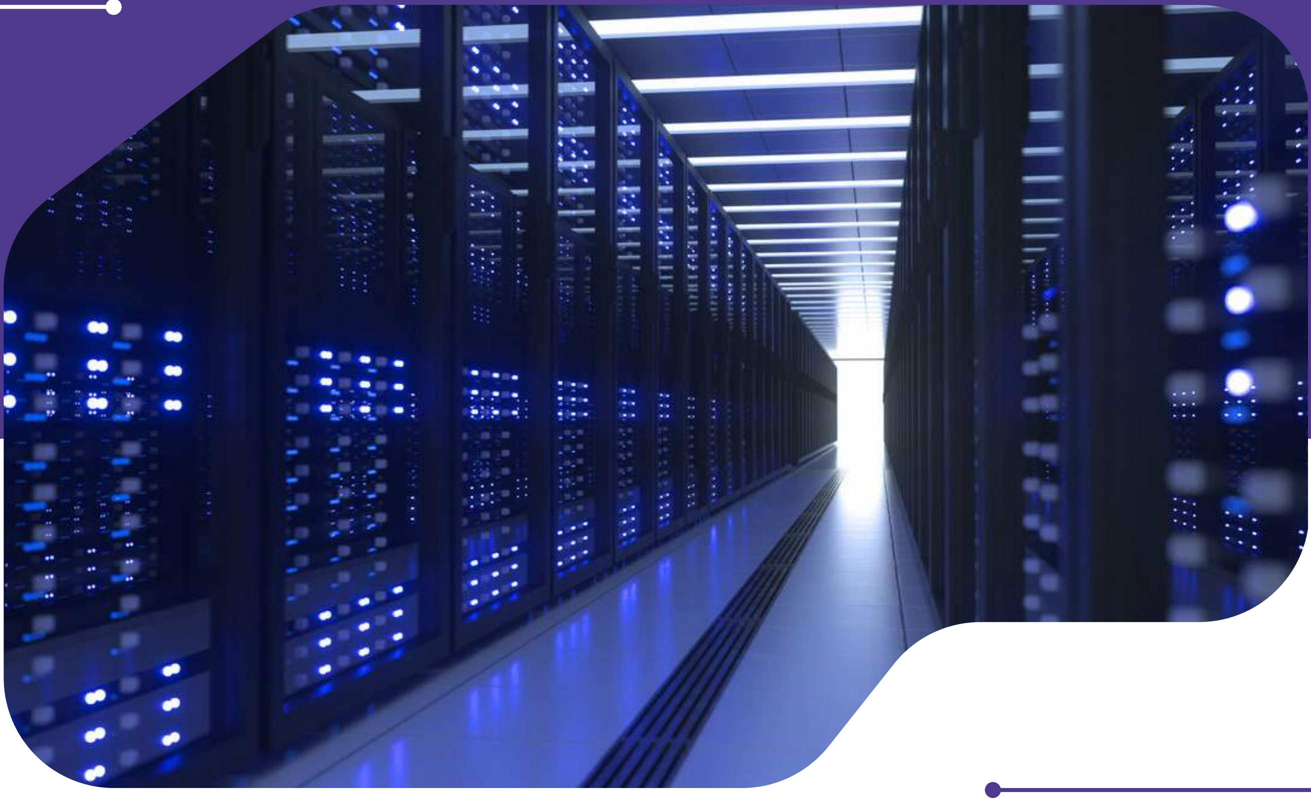
ToneCurve
Technology, LLC



Two-Phase Liquid Cooling Offers a New Model for Cooling Next-Gen Servers

Authored by:
Steven Hill, Independent Technology Analyst

Two-Phase Liquid Cooling Offers a New Model for Cooling Next-Gen Servers

By Steven Hill

The challenge of cooling extremely dense and high-performance infrastructure is a well-known problem, but the recent introduction of two-phase, direct-to-chip liquid cooling options may offer a scalable, efficient, and non-disruptive solution for the long-term challenge of heat management. Removing the heat directly from main sources such as CPUs and GPUs can reduce system heat by as much as 80%, but it takes a two-phase approach to truly make it effective.

Managing Heat is Technology's Perpetual Problem

Heat in the data center is once again becoming a major challenge across the IT industry. This problem has ebbed and flowed for the last two decades as processor density and speed has continued to grow. The production demands of high-performance AI applications and the hardware needed to run them are reaching a new plateau. This demand is driven by turnkey systems like the multi-petaflop capable Nvidia DGX-100H, which features multiple CPUs and GPUs, massive quantities of RAM and SSD storage, and high-speed networking condensed into a single 19"x13"x35" chassis that draws ~10 kW all by itself.

Major industry suppliers have acknowledged this growing trend. In a recent report, Schneider Electric attributed rising chip and rack densities to factors like increased CPU power consumption, wide-scale adoption of high-power GPUs, and, ultimately, low latency production requirements that result in overall increased heat density. Meanwhile, a 2021 report from ASHRAE (the American Society of Heating, Refrigerating and Air-Conditioning Engineers) highlights the growing need for alternatives to air-cooled data centers. In it, ASHRAE covers the multiple and complex challenges of cooling at chip, case, rack, and data center level, while also introducing a clear trend of increasing difficulty in cooling per socket. The chart in Figure 1 below highlights their findings.

"Heat in the data center is once again becoming a major challenge across the IT industry."

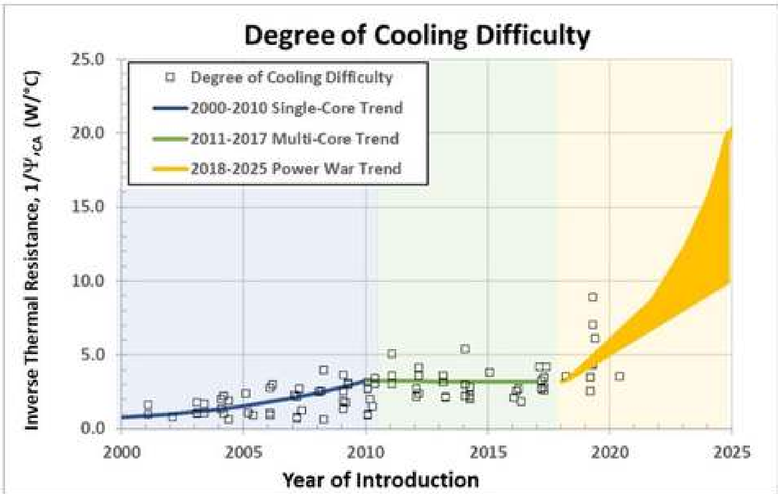


Figure 1. The increasing difficulty of CPU cooling at the socket. (Chart courtesy of ASHRAE.)

This chart illustrates the growing challenge of cooling high-powered chips when factored against power consumption, case temperature, thermal resistance, airflow, and coolant temperatures. It also defines three specific eras of CPU production intent: the single core stage, the multi-core stage, and, most recently, the “power war” stage, in which manufacturers have focused more on attaining maximum performance over other factors in high-end systems.

The State of Air-Cooled Data Centers

For over 50 years, most data center designs have relied on turning large rooms into an analog of a giant refrigerator, drawing heat from the ceiling, running it through a heat exchanger, and then delivering cooled air back to the bottoms of hardware racks to complete the circuit. This concept works just fine while heat loads remain fairly low and consistent from rack to rack, but this is no longer the case for many environments. Now, heat loads in excess of 20kW per rack are rapidly becoming the norm, which present localized extremes that are difficult to manage a generic cold room.

For high-density systems, air cooling remains a nominal option because air is an extremely poor conductor of heat. It’s further limited by the challenge of moving enough air through a given chassis. Ironically, air is actually a key component of many insulation technologies, so simply cooling air and blowing it at a heat source may actually be the least efficient model for cooling. This isn’t new, and many raised-floor operators have been struggling with the need to juggle infrastructure to manage rack heat and power loads for decades. The inexorable physics of thermodynamics may finally be catching up with us, and it seems that air won’t simply be able handle it all anymore.

This is reinforced by the ASHRAE chart in Figure 2 that illustrates a logical transition point to shift from air to liquid cooling. Depending on several factors such as airflow, temperature, and surrounding systems, it will become difficult to cool dense racks of 1U or 2U multi-socketed servers using air alone as CPU power envelopes increase toward ~300W.

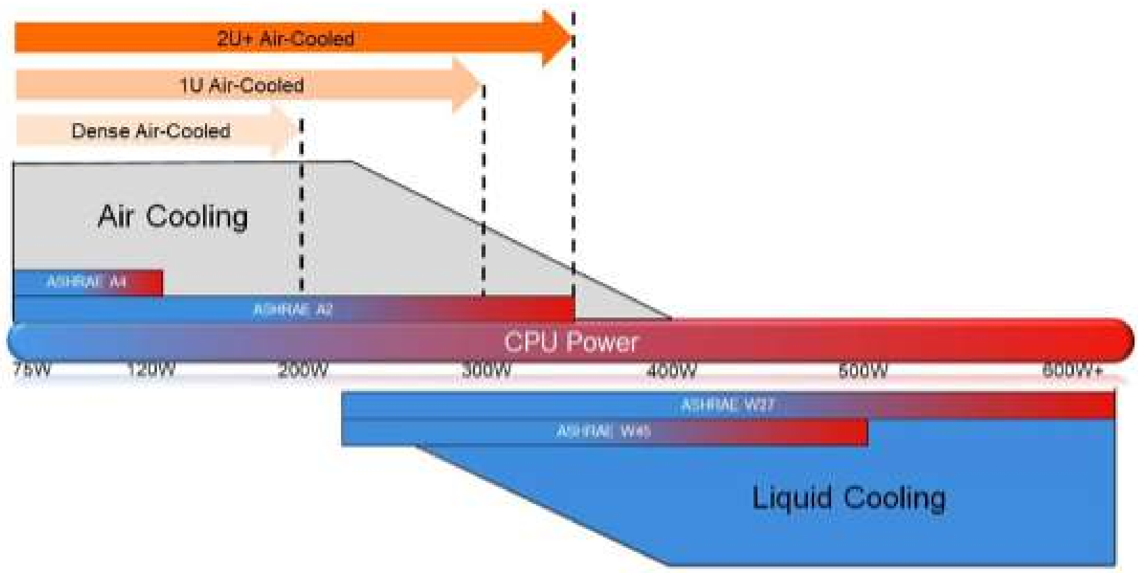


Figure 2. Air-cooling transition point to liquid cooling. (Chart courtesy of ASHRAE.)

A Return to Liquid Cooling

Even though the idea of cooling electronic components with liquid has been around for decades, it has typically been reserved for extremely high-end applications such as supercomputers and mainframes. Water-based cooling immediately comes to mind. However, while water is 23.5 times more efficient in transferring heat than air, it doesn't really work and play well with electronic components when leakage occurs. Additional challenges with water include additional maintenance & service to prevent biofouling and contamination, as well as the need to provide a separate environment dedicated to chilling, pumping, storing, and maintaining water quality. Of course, it's also necessary to monitor and immediately address potential spills.

"...water has joined air-based cooling as an increasingly ineffective coolant for high-compute environments."

Beyond these obvious risks, water has joined air-based cooling as an increasingly ineffective coolant for high-compute environments. Testing has shown that it would require a flow rate of ~2.2 liters per minute to cool a 500W processor, which would substantially increase the amount of pressure needed to force enough liquid to feed the chip cooler. This increase in flow rate can cause strain to the cooling system at higher levels, greatly increasing the risk of leakage. All in all, a more sensible solution would be to use a liquid with a much higher capacity for heat mitigation than single-phase options like water.

An approach that's safe for electronics and better for performance is liquid cooling based on a dielectric fluid that presents no issues with energized components. In immersion cooling, the entire computer is immersed in a large tank of coolant that is recirculated through a secondary cooling system. Single-phase immersion is the most basic form of liquid cooling using oil-based, silicone-based, or engineered fluids that remove the heat from components, but this approach appears to be more popular for specialized applications such as power distribution, battery management, imaging systems, and even data center-adjacent tasks such as bitcoin mining.

More sophisticated is two-phase immersion cooling, where a refrigerant is used as the dielectric fluid and the boiling of that refrigerant provides even greater cooling than possible through the single-phase approach. Refrigerant-based two-phase systems also require a sealed tank system to protect against the escape of gaseous refrigerant. But regardless of the coolant in use, all immersion systems appear to suffer from the same fundamental limitations: all require modifications to the hardware design; all circuitry and cabling must be certified for fluid immersion; and all require a large volume of fluid to operate, often exceeding 800 liters to fill the tank and circulation system. Furthermore, additional infrastructure is needed for filling, storing, and maintaining cooling fluid, hardware is needed to lift components in and out of the tank, and flooring is needed that's capable of bearing a tank's weight, which may exceed the equivalent of a ton per rack.

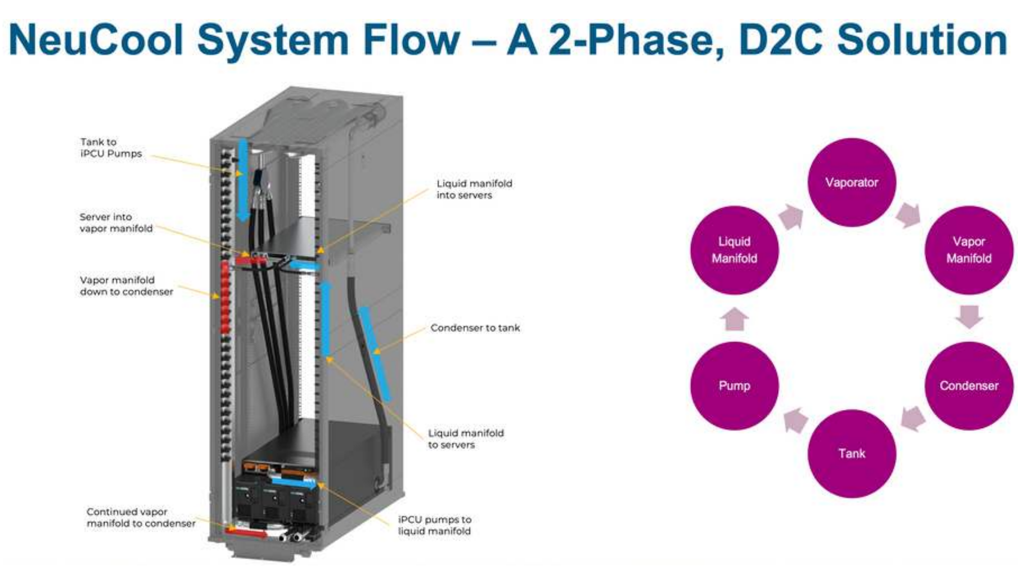
It's easy to understand the initial idea behind immersion cooling, but it's hard to buy into the complexity it adds to achieving the cooling needed for common data center infrastructure. While, yes, it was one of the only options available for cooling extremely high-performance applications more than a decade ago, it doesn't seem to make much sense to submerge an entire system when there are typically only a select few components on the system board that presents serious problems.

Enter: Direct-to-Chip Liquid Cooling

To optimize efficiency of resources, operators have looked increasingly at two-phase, direct-to-chip liquid cooling (DLC) as a strong contender for the next great cooling solution. Two-phase DLC's innovative approach utilizes the thermodynamic principle of latent heat, which refers to the substantial heat energy released or absorbed as a fluid changes state from liquid to gas and back again.

By comparison, single-phase DLC only leverages sensible heat, which refers to energy transferred that causes the fluid's temperature to change (but not its phase). Sensible heat has its limitations—cooling increasingly hotter chips with a single-phase DLC system requires increasing the system's flow rate, using colder facility water, or a combination of the two. Both adjustments have downsides and physical limitations. Meanwhile, for a given facility water temperature and lower system flowrate, two-phase DLC can transfer significantly more heat than single-phase DLC.

Until recently, the challenge has been how to effectively leverage the two-phase process for energy-efficient, reliable, and cost-effective DLC that can operate within the normal constraints of the enterprise data center. However, based on my research, the leader at this point is an emerging company called Accelsius, and their new cooling technology NeuCool. NeuCool uses a refrigerant, selected for its low pressure, low boiling point, and dielectric properties, but its energy-efficient two-phase cycle should not be confused with energy-dense refrigeration systems typically found in data centers. NeuCool uses low-pressure, low-power pumps to circulate refrigerant, unlike refrigeration plants that must compress refrigerant to very high pressures with large, power-hungry compressors.



Fact-Checking the NeuCool Platform

After two decades of serving as a data center technology journalist, analyst, and test lab denizen, I have a finely developed sense for industry hype, as well as a healthy skepticism for “new and revolutionary” technology companies. Any vendor that I’ve briefed with since 2005 would verify that I can be more than a little “challenging” when it comes to digging into their offerings. When presented with the opportunity to do a deep dive on NeuCool, I subsequently warned the team at Accelsius that I’d treat this assignment with a similar dose of skepticism. I’ve pried open enough hardware and climbed enough racks to know what I like, and what I don’t—however, so far, everything I’ve seen in NeuCool has been easy to like.

Here's a quick rundown of my findings so far:

- **Two-phase is for real:** The amount of heat transferred by the transition of refrigerant from liquid to gas is remarkable and can’t be touched by a single-phase approach. Not even close.
- **Low pressure is the way to go:** Because of its greater efficiency, the two-phase approach can operate at a much lower pressure, reduces the potential for leaks, and is easier on the system.
- **Direct to chip makes the most sense:** CPUs and GPUs only take a small amount of real estate on a circuit board, so focused DLC cooling is more efficient, handling ~80% of the heat generated.
- **Modern refrigerants are super effective:** The liquids used by Accelsius have a surprisingly low boiling point, which enables low pressure operation and doesn’t affect room temperature.
- **Their rack-scale approach is smart:** The initial NeuCool system is designed to operate at rack level, which makes it easy to retrofit into existing data centers and can be scaled out as needed.
- **Resilient and easy to maintain:** Cooling pumps and control boards are triple redundant and hot swappable, and the closed loop design only needs ~25 liters of coolant.
- **Minimal server modifications:** The cold plates, or “vaporators” are minimally intrusive and hard-plumbed, and even in the unlikely event of fluid leaking, the refrigerant is dielectric.
- **Compatible with any rack-able server:** the system is certified for most CPU and GPU designs, and work is underway to align with many high-performance server vendors.
- **Future-proof:** the system is currently capable of cooling infrastructure of up to 100kW per rack.

So, you're thinking it can't be perfect, right? What don't I like? At this point, there's literally nothing I dislike. The only caution I'd express is typical—that of dealing with a brand-new company with a brand-new technology. But in the case of a technology like NeuCool, the proof will eventually be in the pudding. Investing in a few racks to deal with trouble spots could be quite cost effective, and it's not like buying into a server, storage, or switching platform, where it often becomes an all-or-nothing proposition for a few buying cycles.

What makes NeuCool even more attractive is its ability to easily slide into any environment with minimal disruption. Think scientific, engineering, industrial, medical, and remote IoT applications where a standalone rack of high-powered gear may be the easiest way to address latency challenges. The possibilities are only limited by your imagination. And you don't have to take my word for it. I have no doubt you'll be hearing about Accelsius's NeuCool system from a number of sources.

I always say that any technology advice that doesn't begin with "well, it depends" should be taken with caution. So, will NeuCool work for you? Well, it depends…but based on everything I've seen so far, I have yet to find a hole in their premise. You should weigh it against the technological needs of your environment, as well as business, financial, and political factors in your enterprise. And, for further background on the intricacies of two-phase liquid cooling, you should definitely check out Dr. Issaam Mudawar's paper, "The Art of Two-Phase Cooling in the Data Center".

It would be easy to joke about how chip technology is "heating up"; however, in reality, the management of heat in the data center has been—and will continue to be—a perpetual challenge for operators. Meanwhile, the liquid cooling industry continues to innovate—and data centers will need to keep pace. As Intel's Dr. Dev Kulkarni said at Data Center World 2024, "it's important to think two or three generations ahead. If you go all out on single-phase only, you might find you need to switch some infrastructure to two-phase technologies within a short period."

Steven Hill is an Independent Technology Analyst and the owner of ToneCurve Technology, LLC.

For over two decades he written hundreds of articles, analyst reports and spoken on evolving data center technologies. In his various positions as both a journalist and an analyst he has focused on a wide range of topics, including: servers, storage, converged infrastructure, data protection, disaster recovery, business continuity, power management and cooling technologies. The opinions expressed in this paper are entirely his own, and may not reflect the position of our sponsor.

This paper has been sponsored by Accelsius. You can find more information on the NeuCool system and how we check all the boxes outlined above at accelsius.com.
